

OBJEKTLAR TO'PLAMINI AVTOMATIK SINFLASH ALGORITMI VA DASTURIY TA'MINOTI

F.Q.Bekmurodov (TATU Samarqand filiali assistenti)

SH.B.Bobojonov (SamDU 2-kurs magistranti)

Ushbu ishda obyektlar to'plamini sinflashda ikkita hol bo'lishi mumkin. Birinchisi obyektlar to'plamini oldindan ma'lum bo'lgan sinflarga bo'lish bo'lsa, ikkinchisi obyektlar to'plamini sinflashda sinflar soni ma'lum bo'lmaydi. Maqolada ikkinchi hol qaralgan bo'lib, unda hosil qilinadigan sinflar soni ma'lum bo'maganda obyektlar qisman pretsedentli prinsip asosida sinflash algoritmi va dasturiy ta'minotini yaratish masalasi qaralgan.

Tabiatda va jamiyatda uchraydigan hodisasalar va jarayonlarda, davlat va boshqaruv organlarida, korxona va tashkilotlarda, xalq xo'jaligining ko'plab sohalarda kiruvchi va chiquvchi ma'lumotlar to'plam shaklida beriladi. Bu sohalarda turli xil kiruvchi va chiquvchi ma'lumotlar to'plamini ma'lum bir mezon asosida sinflash masalasini yechish talab etiladi.

Yuqorida keltirilgan sohalardagi kiruvchi va chiquvchi ma'lumotlarni belgilar deb qarajak, u holda bu ma'lumotlar to'plamini belgilari asosida turli guruxlarga, ya'ni sinflarga ajratish mumkin. Belgilari asosida sinflarga ajratilayotgan hodisalar, jarayonlarni obyektlar deb ataymiz.

Obyektlarni sinflash masalasini amalga oshirishda asosan ikkita yondashuv mavjud. Birinchi yondashuvda obyektlar to'plami o'qituvchi (soha mutaxassisi) tomonidan etalon tanlov shaklida beriladi, ya'ni obyektlar to'plami va obyektlardagi belgilar ro'yxati berilgan bo'lib, obyektlar to'plami oldindan sinflarga bo'lingan holda beriladi. Bu holda sinflash masalasi – etalon tanlov shaklida berilgan, ya'ni oldindan sinflarga bo'lingan obyektlar to'plamini izlab topilgan muhim belgilar yordamida o'zlari joylashgan sinflarga qaytadan sinflashdan iborat.

Ikkinchi yondashuvda obyektlar to'plami o'qituvchi (soha mutaxassisi) tomonidan etalon tanlov shaklida berilmaydi, ya'ni obyektlar to'plami va obyektlardagi belgilar ro'yxati berilgan bo'lib, obyektlarning sinflarda joylashganligi haqidagi ma'lumotlar berilmaydi. Bu holda sinflash masalasi – obyektlar to'plamini sinflarga bo'lishdan iborat. Obyektlar to'plamini sinflarga bo'lish ma'lum bir mezon asosida amalga oshiriladi.

Obyektlar to'plamini sinflashda ikkita hol bo'lishi mumkin. Birinchisi obyektlar to'plamini oldindan ma'lum bo'lgan sinflarga bo'lish bo'lsa, ikkinchisi obyektlar to'plamini sinflashda sinflar soni ma'lum bo'lmaydi. Ikkinchi holda hosil qilinadigan sinflar soni ma'lum bo'maganda obyektlar to'plami avtomatik ravishda sinflarga yoki guruxlarga ajratiladi, bunda sinflar yoki guruxlar soni obyektlar to'plamini sinflash jarayonida hosil bo'ladi.

Maqolada ikkinchi hol qaralgan bo'lib, unda hosil qilinadigan sinflar soni ma'lum bo'maganda obyektlar qisman pretsedentli prinsip asosida sinflash algoritmi va dasturiy ta'minotini yaratish masalasi qaralgan.

Sinflash - bu obyektlarni ularning umumiy xossalari asosida sinflarga bo'lish.

Timsollarni sinflashning asosiy vazifasi - bu obyektlarning bir-biriga mosligini o'rganishda ishlatiladigan, ya'ni ularning o'xshashligini aniqlaydigan hal qiluvchi qoidalarni topishdan iborat. Ikkita timsol bir-biriga o'xshash deyiladi, agarda ular o'rtasida moslik mavjud bo'lsa.

Masalaning qo'yilishi standart shaklda, ya'ni $S_1, S_1, \dots, S_m$ obyektlar to'plamidan iborat T_{nm} tanlov (n - obyektlardagi belgilar soni, m - obyektlar soni) berilgan bo'lsin. T_{nm} tanlovdagi j - obyektning $S_j = a_{j1}, a_{j2}, \dots, a_{jm} (j = 1, m)$ obyektlardagi belgilar alfavitini berish mumkin.

T_{nm} tanlovdagi $S_1, S_1, \dots, S_m$ obyektlarning belgilar alfaviti binar, uzukliksiz sonlar, kesmadagi nuqtalar va nominal (sifatli ko'rsatkichlar) belgilardan iborat bo'lishi mumkin.

T_{nm} tanlovdagi $S_1, S_1, \dots, S_m$ obyektlarni o'qituvchisiz va o'zini-o'zi o'rganish jarayonida sinflarga ajratishdan, ya'ni T_{nm} dan T_{nml} tanlovni hosil qilishdan iborat, bu erda l - hosil qilinishi kerak bo'lgan sinflar soni. Obyektlarni sinflarga ajratishda optimal yechimga ega bo'lish uchun sinflarga ajratuvchi R funksiya minimal $I(R) = \min$ qiymatga ega bo'lishi kerak.

Obyektlar to'plamini qisman presedentli prinsip asosida sinflash algoritmini keltiramiz[2, 7]. Bu algoritm $S = \{S_1, S_2, \dots, S_m\}$ obyektlar to'plamidan Z tayanch obyektlarga nisbatan o'xshash obyektlarni topishga asoslangan. Ushbu algoritm $S = \{S_1, S_2, \dots, S_m\}$ to'plamdan olingan Z tayanch obyektlar bilan ushbu to'plamdagi obyektlarni kema-ket taqqoslash natijasida o'xshashlarini bitta sinfga birlashtiradi.

Algoritmi quyidagi qadamlardan iborat:

1. Obyektlar to'plami $S = \{S_1, S_2, \dots, S_m\}$ qo'rinishda hosil qilinadi.
2. $S = \{S_1, S_2, \dots, S_m\}$ to'plamdan ixtiyoriy S_i obyekt tayanch obyekt sifatida tayinlanadi, ya'ni $Z = \forall S_i$.
3. T_{nm} tanlovdagi obyektlarning belgilar shkalasiga mos $d_1, d_2, \dots, d_k$ taqqoslash qoidalari tanlanadi.
4. $S = \{S_1, S_2, \dots, S_m\}$ to'plamda obyektlarning t - tartibli joylashishi aniqlanadi, ya'ni

$$S_1^t, S_2^t, \dots, S_m^t \quad (1)$$

5. T_{nm} tanlovda n o'lchovga ega bo'lgan boshlang'ich boshqaruvchi vektori $\tilde{b}^0 = \underset{n}{1^1 1^2 \dots 1^k}$ tanlanadi.

6. (1) formulada keltirilgan obyektlardan birinchi uchragan S_1^t obyekt olinadi va $\tilde{b}_1$ boshqaruvchi mantiqiy vektorining qiymati hisoblanadi:

$$b_{1\alpha} = d_\alpha(Z, S_1^t), \alpha = \overline{1, k}.$$

Natijada $\tilde{b} = b_1, b_2, \dots, b_k$ boshqaruvchi vektoriga ega bo'lgan shaxsiy to'plam $D^1(Z, S_1^t)$ hosil bo'ladi.

7. Hosil qilingan $\tilde{b}_1$ boshqaruvchi vektorning qiymati oldingi qadamda hosil qilingan $\tilde{b}^0$ boshqaruvchi vektorning qiymati bilan taqqoslanadi. Agar $\tilde{b}^0 = \tilde{b}^1$ bo'lsa, u holda keyingi qadamga o'tiladi. Agar $\tilde{b}^1 = 0, 0, \dots, 0, 0$, bo'lsa, u holda keyingi qadamga $\tilde{b}^1 = \tilde{b}^0$ boshqaruvchi vektor bilan o'tiladi.

8. $j = j + 1$. Agar $j \leq m$ bo'lsa, u holda 9-qadamga, aks holda 10- qadamga o'tadi.

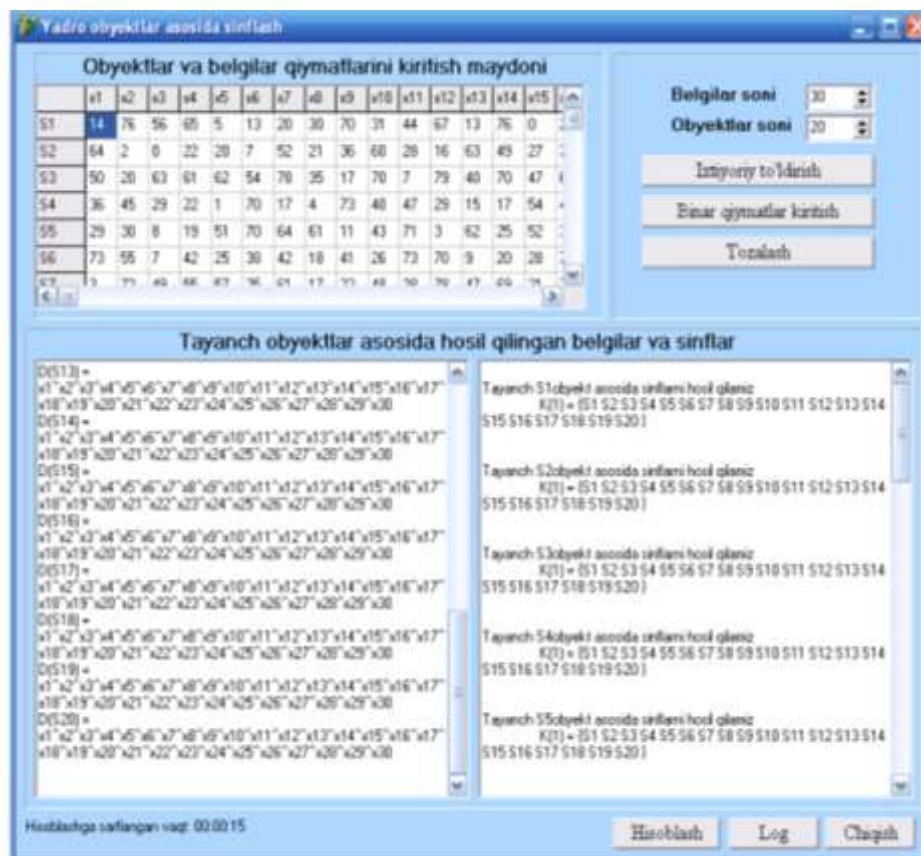
9. Agar $\tilde{b}^{j+1} = \tilde{b}^j$ bo'lsa, u holda Z tayanch obyektini navbatdagi obyekt bilan solishtirish uchun $\tilde{b}^j$ vektordan foydalaniladi. Agarda $\tilde{b}^{j+1} \neq \tilde{b}^j$ bo'lsa, u holda Z tayanch obyektini navbatdagi obyekt bilan solishtirish uchun $\tilde{b}^{j+1}$ vektordan foydalaniladi.

10. Z tayanch obyekt orqali boshqa sinfnii hosil qilish uchun 4 - qadamda $S_1^t, S_2^t, \dots, S_m^t$ obyektlarning joylashish tartibi o'zgartiriladi va algoritm 4 – qadamga o'tadi. Obyektlarning joylashish tartiblari sonini S deb olsak, u holda ushbu protsedura $t \leq C$ shart bajarilguncha davom ettiriladi. Agarda $t > C$ bo'lsa, u holda 12- qadamga o'tiladi.

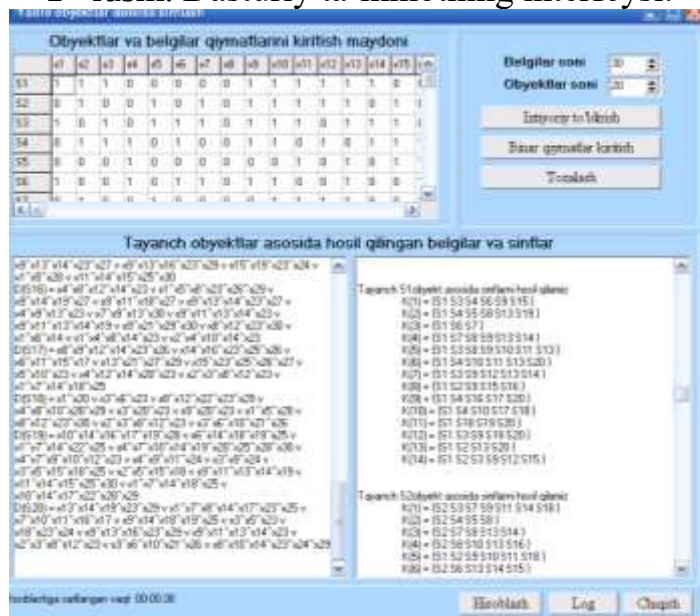
11. $i = i + 1$. Agar $i \leq m$ bo'lsa, u holda $S = \{S_1, S_2, \dots, S_m\}$ to'plamda Z tayanch obyekt o'zgartiriladi, ya'ni $Z = \forall S_{i+1}$ tayinlanadi va 2-qadamga o'tiladi. Aks holda 12-qadamga o'tiladi.

12. $S = \{S_1, S_2, \dots, S_m\}$ to'plamda m ta $Z_1, Z_2, \dots, Z_m$ tayanch obyektlar orqali shaxsiy to'plamlar hosil bo'ladi va har bir shaxsiy to'plamlar o'zining Z tayanch obyektiga va unga o'xshash obyektlar to'plamiga(soniga) ega bo'ladi. Ushbu obyektlar to'plami alohida sinflarni tashkil qiladi.

Ushbu algoritm asosida dasturiy ta'minot yaratildi(1-2-rasmlar).



1 –rasm. Dasturiy ta'minotning interfeysi.



2 –rasm. Dasturiy ta'minotning natijalari.

Dastur obyektlar soni 20 ta belgilar soni 30 ta bo'lganda sinovdan o'tkazildi. Dastur natigasida har bir tayanch obyektga nisbatan bir nechta sinflar hosil qilindi. Masalan S₁ tayanch obyektga nisbatan 14 ta sinf hosil qilindi (3 –rasm)

Tayanch S_1 obyekt asosida sinflarni hosil qilamiz

$K(1) = \{S_1 S_3 S_4 S_6 S_9 S_{15}\}$
 $K(2) = \{S_1 S_4 S_5 S_8 S_{13} S_{19}\}$
 $K(3) = \{S_1 S_6 S_7\}$
 $K(4) = \{S_1 S_7 S_8 S_9 S_{13} S_{14}\}$
 $K(5) = \{S_1 S_3 S_8 S_9 S_{10} S_{11} S_{13}\}$
 $K(6) = \{S_1 S_4 S_{10} S_{11} S_{13} S_{20}\}$
 $K(7) = \{S_1 S_3 S_9 S_{12} S_{13} S_{14}\}$
 $K(8) = \{S_1 S_2 S_9 S_{15} S_{16}\}$
 $K(9) = \{S_1 S_4 S_{16} S_{17} S_{20}\}$
 $K(10) = \{S_1 S_4 S_{10} S_{17} S_{18}\}$
 $K(11) = \{S_1 S_{18} S_{19} S_{20}\}$
 $K(12) = \{S_1 S_3 S_9 S_{19} S_{20}\}$
 $K(13) = \{S_1 S_2 S_{13} S_{20}\}$
 $K(14) = \{S_1 S_2 S_3 S_9 S_{12} S_{15}\}$

3 –rasm. S_1 tayanch obyektga nisbatan hosil qilingan sinflar.

Dasturiy paketni ma'lumotlar obyektlar to'plami ko'rinishida berilgan amaliy masalalarni, ya'ni kvartiralarni ichki va tashqi xarakteristika(yashash maydoni, qurilish materiallari, joylashi, qavat, metrodan uzoqligi va h.k)lari asosida narxi bo'yicha sinflash, avtomobil ishlab chiqarish bozorini sinflash, ichimliklarni ximiyaviy taxlili bo'yicha sinflash, aholining yoshi bo'yicha mamlakatlarni sinflash, radiyosignallarni sinflash, avtomobillar, imzolar, yuzlar, qo'llar tasvirlarini sinflash, qo'lyozma raqam, harf va simvollarini sinflash, tibbiyoda bemorlarni belgilariga asosan kasallik turlariga ajratish, qishloq xo'jaligida mevalarni belgilariga asosan turlarga ajratish, geologiya sohasida foydali qazilmalarni ximiyaviy taxlili bo'yicha turlarga ajratish va boshqa sohalarda qo'llash mumkin.

Foydalanilgan adabiyotlar

1. В.И. Васильев. Распознающие системы. Киев.:Наукова Думка.1986.- 415 с.
2. Фор А. Восприятия распознавания образов. М.1989. 271 с.